

(11) Publication Number: 1994-0022312

(43) Publication Date: October 20, 1994

(21) Patent Application Number: 10-1993-0003922

(22) Filing Date: March 15, 1993

(71) Applicant: Yeongtaek KIM, Hyoseok LEE, Gwangseop SIM, Boryeon KIM, Seonghee YOON, Seunghyeon YANG, Seongdong KIM, Jaewon LEE, Yuseop KIM

(72) Inventors: Yeongtaek KIM, Hyoseok LEE, Gwangseop SIM, Boryeon KIM, Seonghee YOON, Seunghyeon YANG, Seongdong KIM, Jaewon LEE, Yuseop KIM

(54) Title: MACHINE TRANSLATION APPARATUS AND METHOD THEREOF

Abstract:

The present invention relates to a machine translation apparatus and method thereof for translating a original text described in a first natural language into a translated text described in a second natural language, and more particularly to a machine translation apparatus and method thereof using a procedure parsing and converting sentence according to phrases and language. To obtain a proper Korean sentence from the result of parsing syntax of an English sentence which is a machine transition language input sentence between two conventional languages, a complex and repeating structure conversion procedure of several steps and a word substitution procedure are required. So, the translation process becomes complex, and the translated contents are not exact. Accordingly, exact translation results can be obtained by using phrase unit transition based on phrase and compound. Various expression patterns of a primitive language from which excellent translation is difficult to achieve are classified into phrases, and the translation dictionary that has correspondence of the various expression patterns of the primitive language to translation phrase of a target language is provided to enhance the translation exactness.

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(51) Int. Cl.
G06F 15/38

(11) 공개번호
(43) 공개일자

특1994-0022312
1994년10월20일

(21) 출원번호 특1993-0003922

(22) 출원일자 1993년03월15일

(71) 출원인 김영택
대한민국
서울특별시 서초구 반포본동 반포아파트 90-406
이호석
대한민국
경기도 광명시 하안3동 주공아파트 703-403
심광섭
대한민국
서울특별시 성북구 석관동 43-36
김보령
대한민국
서울특별시 관악구 신림9동 244-8
윤성희
대한민국
서울특별시 서초구 서초4동 상풍아파트 3동 301호
양승현
대한민국
서울특별시 관악구 신림본동 11-8 웨미리 나동 B101
김성동
대한민국
서울특별시 종로구 혜화동 126-5호 8동 1반
이재원
대한민국
서울특별시 송파구 석촌동 10-10
김유섭
대한민국
서울특별시 서초구 반포본동 반포아파트 90-406
(72) 발명자 김영택
대한민국
서울특별시 서초구 반포본동 반포아파트 90-406
이호석
대한민국
경기도 광명시 하안3동 주공아파트 703-403
심광섭
대한민국
서울특별시 성북구 석관동 43-36
김보령
대한민국
서울특별시 관악구 신림9동 244-8
윤성희
대한민국
서울특별시 서초구 서초4동 상풍아파트 3동 301호
양승현
대한민국
서울특별시 관악구 신림본동 11-8 웨미리 나동 B101
김성동
대한민국
서울특별시 종로구 혜화동 126-5호 8동 1반
이재원
대한민국

서울특별시 송파구 석촌동 10-10
 김유성
 대한민국
 서울특별시 서초구 반포본동 반포아파트 90-406
 이영필
 김용식
 이윤민
 있음
 기계번역장치 및 방법

(74) 대리인

(77) 심사청구

(54) 출원명

요약

본 발명은 제1언어로 기술된 자연언어(원문)를 번역처리하여 제2언어로 기술된 자연언어(번역문)로 출력하도록 하는 기계번역처리장치 및 그 방법에 관한 것으로, 특히 속어와 언어에 의한 문장의 분석과 변환 과정을 이용한 기계번역장치 및 그 방법에 관한 것으로, 종래의 두 언어간의 기계번역이 입력문장인 영어문장에 대한 구문분석결과로부터 적절한 한국어 문장을 얻기 위해서는 여러 단계의 복잡하고 반복적인 구조변환과정과 어휘대치과정을 거침으로 인하여 번역이 복잡해지고 그 내용이 정확하지 않은 문제점이 있어 속어와 언어를 중심으로 하는 구절단위의 번역에 의하여 정확한 번역결과를 얻을 수 있으며, 또한 우수한 번역문을 얻기 곤란한 원시언어의 다양한 표현 형태들을 속어로 구분하고 이에 대한 번역구의 대응을 갖춘 번역 사전을 제공하여 번역의 정확성을 높일 수 있는 효과가 있다.

대표도

도1

영세서

[발명의 명칭]

기계번역장치 및 방법

[도면의 간단한 설명]

제1도는 종래의 기계번역장치의 개략블럭도,

제2도는 본 발명에 따른 기계번역장치의 블럭도,

제7도는 본 발명에 따른 기계번역장치의 전역구조화부에서 이루어지는 전역구조분석과정을 일예를 도시한 도이다.

본 내용은 요부공개 건이므로 전문 내용을 수록하지 않았음.

(57) 청구의 범위

청구항 1.

제1언어로 된 각종 품사의 단어를 포함하는 어휘사전; 상기 어휘사전을 참조하여 제1언어의 입력 문장을 어휘별로 분석하는 어휘분석수단; 분석된 어휘자료 및 번역자료를 기록하는 활성화트; 제1언어를 제2언어로 번역하기 위한 어휘자료를 포함하는 번역사전; 제1언어의 문법규칙을 포함하는 제1언어규칙부; 상기 번역사전에서 제공하는 속어패턴들을 상기 활성화트에 기록된 데이터들과의 매칭과 속어변수에 대한 지역적 탐색에 의하여 속어를 인식하여 이를 비활성제지로서 활성화트에 기록하는 속어인식수단; 상기 번역사전을 이용하여 상기 제1언어규칙부에 규정되어 있는 문법규칙에 따라 상기 활성화트에 기록된 데이터로 문장의 분석을 전역적으로 행하기 위한 전역구조화수단; 전역적으로 분석된 데이터를 제2언어로 변환하기 위한 어휘정보를 제공하는 생성사전; 제2언어의 문법규칙을 포함하는 생성규칙부; 및 상기 생성사전을 이용하여 상기 생성규칙부에 규정되어 있는 문법규칙에 따라 제2언어문장을 생성하기 위한 제2언어무장생성부를 포함하는 것을 특징으로 하는 기계번역 장치.

청구항 2.

제1항에 있어서, 상기 기계번역장치는 상기 전역구조화수단에 의하여 처리된 데이터로부터 상기 번역사전을 참조하여 언어쌍을 검색하여 제1언어의 어휘를 번역하기 위한 언어처리수단을 더 포함하는 것을 특징으로 하는 기계번역장치.

청구항 3.

제1항 또는 제2항에 있어서, 상기 제1언어는 영어이고 제2언어는 한국어인 것을 특징으로 하는 기계번역장치.

청구항 4.

제1항 또는 제2항에 있어서, 상기 어휘사전은 단어들의 기본형태와 불규칙 변형된 형태만을 포함하는 것을 특징으로 하는 기계번역장치.

청구항 5.

제1항 또는 제2항에 있어서, 상기 활성화트는 상향식 제어와 하향식제어를 결합한 방식을 사용하는 것을 특징으로 하는 기계번역장치.

청구항 6.

제1항 또는 제2항에 있어서, 상기 번역사전은 기록되어 있는 각 단어에 대하여 속어구, 언어상 및 기본형에 대한 정보를 포함하는 것을 특징으로 하는 기계번역장치.

청구항 7.

제6항에 있어서, 상기 번역사전은 사전을 고속으로 검색할 수 있도록 인덱스를 구성하고 검색기능을 수행하는 사전제어부를 포함하는 것을 특징으로 하는 기계번역방법.

청구항 8.

제1항 또는 제2항에 있어서, 상기 활성차트내에 등록된 비활성어지는 상기 전역구조화부에서의 전역구조화과정에서 보류되는 것을 특징으로 하는 기계번역장치.

청구항 9.

제1항 또는 제2항에 있어서, 상기 번역사전은 기본사전과 분야사전으로 구성되며, 동일한 어휘에 대하여 분야사전이 기본사전보다 우선적으로 참조되는 것을 특징으로 하는 기계번역장치.

청구항 10.

제1항 또는 제2항에 있어서, 상기 규칙부는 조건을 가지는 구절구조규칙들의 집합 및 특정 규칙을 선택하기 위한 규칙선택수단을 포함하는 것을 특징으로 하는 기계번역장치.

청구항 11.

제2항 또는 제2항에 있어서, 상기 언어처리수단에 의하여 언어쌍으로써 처리되는 대상은 타동사와 그 목적어의 쌍, 자동사와 주어의 쌍, 및 형용사와 수식되는 명사의 쌍에 대하여 적용되는 것을 특징으로 하는 기계번역방법.

청구항 12.

제1항 또는 제2항에 있어서, 상기 제2언어문장생성부는 복수어 표현을 분석하여 구조를 정규화하는 복수어분석수단; 상기 복수어분석수단에서 처리된 구조를 생성규칙에 따라 통사전변형 및 어순조정을 행하는 구문생성수단; 및 상기 구문생성수단에 의하여 생성된 각 구문들에 대한 형태소의 첨가 및 합성을 행하여 완성된 제2언어의 문장을 생성하는 형태소생성수단을 포함하는 것을 특징으로 하는 기계번역장치.

청구항 13.

제1언어로 이루어진 원시언어 입력문장을 제2언어인 목적언어로 번역하기 위하여 제1언어에 대한 어휘사전, 제2언어의 번역사전, 번역을 위하여 사용되는 규칙부 등을 포함하는 제1언어를 제2언어로 번역하는 기계번역방법에 있어서, 상기 번역방법이 입력된 제1언어의 문장을 어휘사전을 이용하여 각 어휘별로 분석하는 어휘분석단계; 분석된 문장의 단어들을 전역자료구조인 활성차트의 초기상태를 이루도록 하는 자료구조 초기화단계; 상기 초기화된 자료구조에서 번역사전을 이용하여 문법규칙에 따라 특정 단어를 키워드로 하여 문장의 특정 지점을 시작점으로 하는 지역적 탐색과정을 통하여 요구되는 숙어변수를 대치할 구조를 검색하여 인식된 숙어를 비활성어지로서 등록하는 숙어인식단계; 상기 숙어인식단계의 종료후 문법규칙에 따라 인식된 숙어를 중심으로 각 활성어지에 대한 규칙의 선택 및 적용으로 새로운 예지를 추가하는 전역적 탐색과정을 통하여 문장전체를 결합하는 구조를 만드는 전역구조화단계; 상기 구조화단계후 분석된 구조중에서 번역되지 않은 나머지 단어들에 대해서는 디폴트 번역어휘를 사용하여 번역을 행하는 디폴트어휘번역단계; 및 상기 전역구조화단계에서의 결과를 제2언어의 생성규칙과 생성사전을 참조하여 제2언어의 문장을 생성하는 제2언어생성단계를 포함하는 것을 특징으로 하는 기계번역방법.

청구항 14.

제13항에 있어서, 상기 전역구조화단계후에 번역사전을 이용하여 상기의 구조내의 어휘에 대한 언어정보를 추출하여 번역처리하는 언어처리단계를 더 포함하는 것을 특징으로 하는 기계번역방법.

청구항 15.

제13항 또는 제14항에 있어서, 상기 제2언어 생성단계는 상기 생성단계이전까지의 처리된 구조에서 복수어 표현을 분석하여 구조를 정규화하는 복수어분석단계; 상기 분석단계에서 처리된 구조를 생성규칙에 따라 통사전변형 및 어순조정을 행하는 구문생성단계; 및 상기 구문생성단계에 의하여 생성된 구문들에 대한 형태소의 첨가 및 합성을 행하여 완성된 제2언어의 문장을 생성하는 형태소생성단계를 포함하는 것을 특징으로 하는 기계번역방법.

청구항 16.

제13항 또는 제14항에 있어서, 상기 제1언어는 영어이고, 제2언어는 한국어인 것을 특징으로 하는 기계번역방법.

청구항 17.

제13항 또는 제14항에 있어서, 상기 숙어인식단계에서는 이루어지는 숙어의 유형은 고정된 단어들의 열로 이루어지는 단순숙어와 고정단어로 시작되며, 고정단어들과 한개 또는 두개의 숙어변수(메타심볼)를 가지는 메타숙어인 것을 특징으로 하는 기계번역방법.

청구항 18.

제13항 또는 제14항에 있어서, 번역사전은 기본사전과 분야사전으로 나누어져 있으며 번역에 있어서는 분야사전이 기본사전에 우선하여 참조되는 것을 특징으로 하는 기계번역방법.

청구항 19.

제14항에 있어서, 언어처리는 타동사와 그 목적어의 쌍, 자동사와 주어의 쌍, 및 형용사와 수식되는 명사의 쌍에 대하여 적용되는 것을 특징으로 하는 기계번역방법.

청구항 20.

제13항 또는 제14항에 있어서, 숙어인식단계에서 얻어지는 비활성어지는 그 예지의 시작점과 끝점사이에 존재하는 점들이 전역구조화과정에서 보류되는 것을 특징으로 하는 기계번역방법.

청구항 21.

제13항 또는 제14항에 있어서, 숙어인식단계에서 얻어지는 비활성어지는 그 예지의 시작점과 끝점사이에 존재하는 점들이 전역구조화과정에서 보류되는 것을 특징으로 하는 기계번역방법.

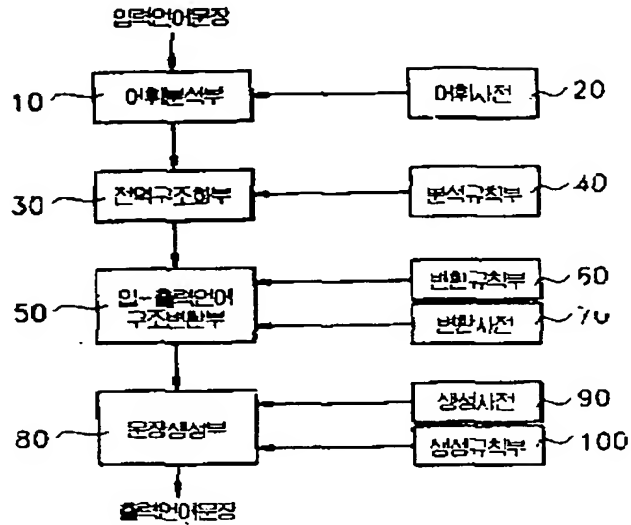
청구항 22.

"21항에 있어서, 상기 지역적 탐색은 속어변수에 대한 예측을 도입하여 전체 규칙집합중에서 부분집합만이 적용되는 것을 특징으로 하는 기계 번역방법.

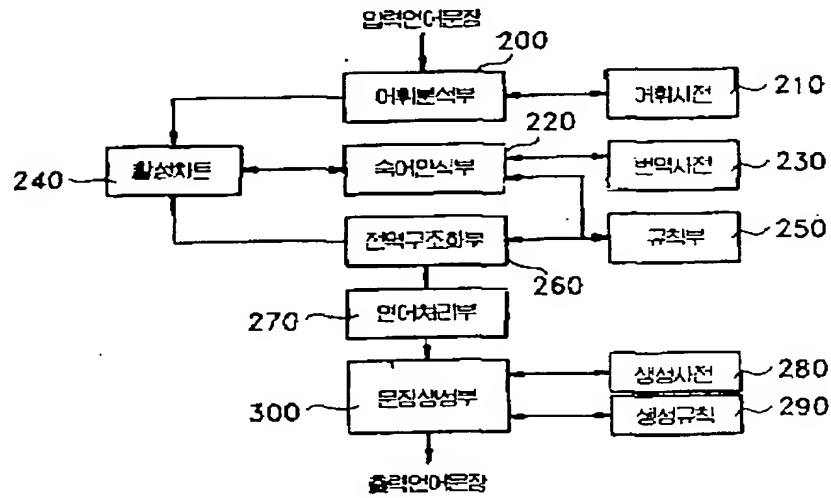
※ 참고사항 : 최초출원 내용에 의하여 공개하는 것임.

도면

도면 1



도면 2



도면 7

"It prevents you from accessing two records at the same time."

(2) (1)

(3)

(4)

(5)

(6)